

# Interference Alignment with Analog Channel State Feedback

Omar El Ayach and Robert W. Heath, Jr.

## Abstract

Interference alignment (IA) is a multiplexing gain optimal transmission strategy for the interference channel. While the achieved sum rate with IA is much higher than previously thought possible, the improvement often comes at the cost of requiring network channel state information at the transmitters. This can be achieved by explicit feedback, a flexible yet potentially costly approach that incurs large overhead. In this paper we propose analog feedback as an alternative to limited feedback or reciprocity based alignment. We show that the full multiplexing gain observed with perfect channel knowledge is preserved by analog feedback and that the mean loss in sum rate is bounded by a constant when signal-to-noise ratio is comparable in both forward and feedback channels. When signal-to-noise ratios are not quite symmetric, a fraction of the multiplexing gain is achieved. We consider the overhead of training and feedback and use this framework to optimize the system's effective throughput. We present simulation results to demonstrate the performance of IA with analog feedback, verify our theoretical analysis, and extend our conclusions on optimal training and feedback length.

## I. INTRODUCTION

The interference channel is an information theoretic concept that models wireless networks in which several transmitters simultaneously communicate data to their paired receivers. The traditional approach for communication in such channels was orthogonalization, where resources are split among the various transceiver pairs. Unfortunately, this results in decaying per-user data rates as networks grow. Recent work in information theory, however, has shown that the interference channel can support sum rates that, under certain assumptions, scale linearly with

The authors are with the Wireless Networking and Communications Group, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712 USA (e-mail: {omarayach, rheath}@mail.utexas.edu)

This work was supported by the Office of Naval Research (ONR) under grant N000141010337.

the number of users at high signal-to-noise ratios (SNR). This linear sum rate scaling can be achieved by a transmission strategy known as interference alignment (IA) [1].

Interference alignment is a linear precoding technique that attempts to align interfering signals in time, frequency, or space. In multiple-input multiple-output (MIMO) networks, IA utilizes the spatial dimension offered by multiple antennas for alignment. By aligning interference at all receivers, IA reduces the dimension of interference allowing users to cancel interference via linear techniques and decode their desired signals interference free. In a single-input single-output (SISO) system with infinitely many channel extensions, IA allows users to communicate at approximately half their interference free rate [1]. IA has also been shown to achieve maximum multiplexing gain in a class of extended MIMO channels [2]. While such a general result has not been proven for the constant MIMO channel, antennas have been shown to provide a practical source of dimensionality [3]–[5]. While various constant MIMO IA solutions have been proposed, all solutions assume some form of channel state information (CSI) at the transmitter obtained either implicitly, through channel reciprocity [6], [7], or explicitly through interference pricing [8] or CSI feedback [1], [3], [9]–[12]. Reciprocity, however, does not hold in frequency duplexed systems. Moreover, with time duplexing, reciprocity requires tight calibration of RF devices. Calibration must either be accounted for in the RF hardware [13], or done dynamically using a series of forward and reverse transmissions as in [14] which again incurs overhead. Since reciprocity may not hold, methods to transfer CSI back to the transmitters are important for realizing the sum capacities promised by IA.

CSI feedback potentially incurs large overhead which reduces the effective data rates achieved. Therefore, low overhead feedback strategies that preserve IA's sum rate performance must be used to satisfy the CSI requirement. Unfortunately, work on managing the overhead of IA, however, is rather limited. In [7], the overhead of the proposed reciprocity-based algorithm is numerically analyzed. Others have proposed network design strategies such as partitioning to reduce overhead [15]. Network partitioning, however, only reduces the number of channels that are to be fed back and makes no attempt at improving the underlying feedback strategy itself. A well established method to improve feedback and further reduce overhead is to employ channel state quantization. For example, [16] uses Grassmannian codebooks to directly quantize and feedback the wideband channel coefficients in single antenna systems which employ IA by coding over frequency. This limited feedback approach was later extended in [17] to systems

with multiple antennas at the transmitters, receivers, or both. In both [16] and [17], multiplexing gain is preserved by scaling the number of feedback bits with SNR. The need for such scaling is a characteristic of digital feedback and is indeed in line with earlier results for other multiuser channels [18]. The complexity of quantized feedback, however, increases with codebook size, which is challenging since large Grassmannian codebooks are hard to design and encode.

In this paper we propose using IA with analog feedback [19]–[22]. Instead of quantizing the channel state information, analog feedback directly transmits the channel matrix elements as uncoded quadrature and amplitude modulated symbols. In the proposed strategy, with knowledge of the forward channels, receivers train the reverse links and feedback the forward channel matrices using analog feedback. We show that, under mild assumptions on feedback quality, performing IA on the channel estimates obtained via analog feedback incurs no loss in multiplexing gain [23]. Specifically, using a cooperative analog feedback strategy, we show that when the SNR on both forward and reverse channels is comparable, the loss in sum rate achieved by a linear zero-forcing receiver is upper bounded by a constant which implies the preservation of multiplexing gain. We extend the performance analysis to demonstrate that other analog feedback strategies, which assume no cooperation, perform similarly. Moreover, we show that when such symmetry of SNR is not possible, the system still achieves a fraction of its degrees of freedom. Using the derived bounds on loss in sum rate due to imperfect CSI, we consider the system's effective throughput, accounting for the overhead incurred by training and feedback. We use our constructed framework to optimize training and feedback lengths to find the best operating point on the trade-off curve between overhead and sum rate.

This paper extends the analog feedback framework that was originally proposed for the MISO broadcast channel [19] to the MIMO interference channel, and characterizes its performance when IA is used as a transmission strategy. We show that analog feedback is a viable alternative to the quantization based schemes presented in [16], [17]. As an advantage, analog feedback does not suffer from the same increasing complexity at high SNR as digital feedback, and the required feedback scaling will come naturally in most wireless ad hoc networks. Moreover, analog feedback remains optimal when the number of feedback symbols equals the number of feedback channel uses [21]. We also study the effect of feedback and training overhead on IA's performance, a concept which, with the exception of [7], [15], has been mostly neglected.

Throughout this paper we use the following notation:  $\mathbf{A}$  is a matrix;  $\mathbf{a}$  is a vector;  $a$  is a scalar;

$\mathbf{A}^*$  and  $\mathbf{a}^*$  denote the conjugate transpose of  $\mathbf{A}$  and  $\mathbf{a}$  respectively;  $\|\mathbf{A}\|_F$  is the Frobenius norm of  $\mathbf{A}$ ;  $\|\mathbf{a}\|_p$  is the  $p$ -norm of  $\mathbf{a}$ ;  $\mathbf{I}_N$  is the  $N \times N$  identity matrix;  $\mathbb{C}^N$  is the  $N$ -dimensional complex space;  $\mathcal{CN}(\mathbf{a}, \mathbf{A})$  is a complex Gaussian random vector with mean  $\mathbf{a}$  and covariance matrix  $\mathbf{A}$ ;  $(a_1, \dots, a_k)$  is an ordered set. Other notation is defined when needed.

## II. SYSTEM MODEL AND BACKGROUND

In this section we introduce the MIMO interference channel under consideration. We review the concept of IA with perfect channel state information, with a focus on the equations and properties that will be used in our analysis of analog feedback in later sections.

### A. Interference Channel Model

Consider the narrowband MIMO interference channel shown in Fig. 1. In this channel each of the  $K$  source nodes  $i$  communicates with its sink node  $i$  and interferes with all other sink nodes,  $\ell \neq i$ . We refrain from using the term transmitter-receiver since all nodes will be involved in both transmission and reception in either payload transmission or feedback intervals. For simplicity of exposition, we consider the case of the homogeneous network where each source and sink is equipped with  $N_t$  and  $N_r$  antennas respectively. Thus, source node  $i$  transmits  $d_i \leq \min(N_t, N_r)$  independent spatial streams to its corresponding sink. The results can be readily generalized to a different number of antennas at each node, provided that IA remains feasible [4].

We consider a block fading channel model in which channels are drawn independently across all users and antennas and remain fixed for the interval of interest. We neglect large scale fading effects which can be accounted for at the expense of more involved exposition in Section III-A. We also assume perfect time and frequency synchronization when expressing received baseband signals. Under our assumptions, the received signal at sink node  $i$  can be written as

$$\mathbf{y}_i = \sqrt{\frac{P}{d_i}} \mathbf{H}_{i,i} \mathbf{F}_i \mathbf{s}_i + \sum_{\ell \neq i} \sqrt{\frac{P}{d_\ell}} \mathbf{H}_{i,\ell} \mathbf{F}_\ell \mathbf{s}_\ell + \mathbf{v}_i,$$

where  $\mathbf{y}_i$  is the  $N_r \times 1$  received signal vector,  $\mathbf{H}_{i,\ell}$  is the  $N_r \times N_t$  channel matrix from source  $\ell$  to sink  $i$  with i.i.d  $\mathcal{CN}(0, 1)$  elements,  $\mathbf{F}_i = [\mathbf{f}_i^1, \dots, \mathbf{f}_i^{d_i}]$  is node  $i$ 's  $N_t \times d_i$  unitary precoding matrix,  $\mathbf{s}_i$  is the  $d_i \times 1$  transmitted symbol vector at node  $i$  such that  $\mathbb{E}[\|\mathbf{s}_i\|^2] = d_i$ , and  $\mathbf{v}_i$  is a complex vector of i.i.d complex Gaussian noise with covariance matrix  $\sigma^2 \mathbf{I}_{N_r}$ . We assume equal power allocation since the gain observed from water-filling is negligible at high SNR [24].

We place no reciprocity assumption on the forward and reverse channels, as in a frequency division duplexed system, for example. On this channel, the received signal at *source node*  $i$  is

$$\overleftarrow{\mathbf{y}}_i = \sqrt{\frac{P_f}{N_r}} \mathbf{G}_{i,i} \overleftarrow{\mathbf{x}}_i + \sum_{\ell \neq i} \sqrt{\frac{P_f}{N_r}} \mathbf{G}_{\ell,i} \overleftarrow{\mathbf{x}}_\ell + \nu_i, \quad (1)$$

where  $P_f$  is the transmit power used to transmit pilot and feedback symbols,  $\mathbf{G}_{\ell,i}$  is the  $N_t \times N_r$  reverse channel between sink node  $\ell$  and source node  $i$  with i.i.d  $\mathcal{CN}(0,1)$  elements,  $\overleftarrow{\mathbf{x}}_i$  is the symbol vector with unit norm elements sent by sink  $i$ , and  $\nu_i$  is a complex vector of i.i.d circularly symmetric white Gaussian noise with covariance matrix  $\sigma^2 \mathbf{I}_{N_t}$ .

### B. Interference Alignment

IA for the MIMO interference channel is a linear precoding technique that by potentially coding over infinite channel extensions achieves the channel's degrees of freedom defined as  $\lim_{P \rightarrow \infty} \frac{R_{sum}}{\log_2 P}$ . This result originally assumed that the magnitude of continuously distributed i.i.d channel coefficients is bounded away from zero and infinity [2] to avoid the degenerate cases of equal coefficients or channels equal to zero or infinity. For constant MIMO channels with Rayleigh fading, degenerate cases happen with probability zero and IA can improve the achieved sum rate as shown in [3], [5], [6]. In constant MIMO channels, IA computes the transmit precoders  $\mathbf{F}_i$  to align interference at all receivers in a strict subspace of the received signal space, thus leaving interference free dimensions for the desired signal. While IA is only one of the many precoding strategies for the interference channel [3], [7], [25], some of which marginally outperform it at low SNR [3], a main advantage of IA is that it is analytically tractable. Its complete interference suppression properties make it especially amenable to the study of performance with feedback and imperfect CSI.

While IA can be used with any receiver design, to simplify exposition we consider a per-stream zero-forcing receiver in which sink node  $i$  projects its received signal on to the columns,  $\mathbf{w}_i^m$ , and treats residual interference as noise. Simulations in Section V indicate that the same performance can be expected from a joint signal decoder which again treats interference as noise.

Writing the per stream input-output relation at the output of  $\mathbf{w}_i^m$  gives

$$(\mathbf{w}_i^m)^* \mathbf{y}_i = (\mathbf{w}_i^m)^* \sqrt{\frac{P}{d_i}} \mathbf{H}_{i,i} \mathbf{f}_i^m s_i^m + \sum_{\ell \neq m} (\mathbf{w}_i^m)^* \sqrt{\frac{P}{d_i}} \mathbf{H}_{i,i} \mathbf{f}_i^\ell s_i^\ell + \sum_{k \neq i} \sum_{\ell=1}^{d_k} (\mathbf{w}_i^m)^* \sqrt{\frac{P}{d_k}} \mathbf{H}_{i,k} \mathbf{f}_k^\ell s_k^\ell + (\mathbf{w}_i^m)^* \mathbf{v}_i,$$

for  $m \in \{1, \dots, d_i\}$  and  $i \in \{1, \dots, K\}$ , where  $\|\mathbf{w}_i^m\|^2 = 1$ . At the output of these linear receivers  $\mathbf{w}_i^m$ , the conditions for perfect interference alignment can be restated as [6]

$$(\mathbf{w}_i^m)^* \mathbf{H}_{i,k} \mathbf{f}_k^\ell = 0, \quad \forall (k, \ell) \neq (i, m) \quad (2)$$

$$|(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m| \geq c > 0, \quad \forall i, m \quad (3)$$

where IA is guaranteed by the first condition, and the second is satisfied with high probability.

The sum rate achieved by such a linear zero-forcing receiver, assuming Gaussian input signals and treating leakage interference as noise, is

$$R_{sum} = \sum_{i=1}^K \sum_{m=1}^{d_i} \log_2 \left( 1 + \frac{\frac{P}{d_i} |(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m|^2}{\mathcal{I}_{i,m} + \sigma^2} \right), \quad (4)$$

where  $\mathcal{I}_{i,m}$  is the total inter-stream and inter-user interference given by [18], [21]

$$\mathcal{I}_{i,m} = \sum_{(k,\ell) \neq (i,m)} \frac{P}{d_k} |(\mathbf{w}_i^m)^* \mathbf{H}_{i,k} \mathbf{f}_k^\ell|^2.$$

In the presence of perfect channel knowledge, and for an achievable degree of freedom vector  $\mathbf{d} = [d_1, d_2, \dots, d_K]$ , equations (2) and (3), are satisfied and thus  $\mathcal{I}_{i,m} = 0$ . This gives

$$\lim_{P \rightarrow \infty} \frac{R_{sum}}{\log_2 P} = \lim_{P \rightarrow \infty} \frac{\sum_{i,m} \log_2 \left( 1 + \frac{\frac{P}{d_i} |(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m|^2}{\sigma^2} \right)}{\log_2 P} = \sum_{i=1}^K d_i \leq \min(N_t, N_r) \frac{R}{R+1} K,$$

in the case where  $R = \frac{\max(N_t, N_r)}{\min(N_t, N_r)}$  is an integer such that  $K > R$  [2]. The full characterization of the extended MIMO interference channel's degrees of freedom is provided in [2].

It is not immediately clear, however, if the same sum rate scaling behavior can be expected from a network with only imperfect knowledge of the channel derived from analog feedback. Results on single user MIMO generally prove an acceptable constant rate loss due to imperfect CSI [26]. In most multi-user scenarios, however, the cost of imperfect CSI may be much higher, potentially resulting in the loss of the channel's multiplexing gain [18] which saturates performance at high SNR [27]. In Section III, we show that such performance can be expected from a realistic IA system with analog feedback, provided that the quality of channel knowledge scales sufficiently with transmit power, or effectively the forward channel's SNR. The result is based on the necessary accuracy of CSI which must scale with SNR as shown in [28]. This is

similar to the results presented in [16] and [17] for IA with digital feedback where CSI quality is scaled by controlling codebook size and feedback bits.

### III. INTERFERENCE ALIGNMENT WITH ANALOG FEEDBACK

In this section we propose a feedback and transmission strategy based on analog feedback and naive IA, which uses the estimated channels as if they were the true propagation channels.

#### A. Analog Feedback

To feedback the forward channel matrices  $\mathbf{H}_{i,\ell}$  reliably across the feedback channel given in (1), we propose dividing the feedback stage into two main phases [19]. The details of the analog feedback strategy considered are summarized in Fig. 1. First, each source independently learns its reverse channels. Second, the forward channels are fed back and estimated. We neglect forward channel training and estimation and, thus, assume they have been estimated perfectly. This is since imperfect CSI at the receiver adversely affects all feedback schemes and is not exclusive to analog feedback. In fact, the analysis of forward channel estimation error parallels the analysis in Section III-A1 and will simply add an extra error term that decays with power as needed in Section III-B. Therefore, similar results can be readily shown.

1) *Reverse Link Training:* To learn all reverse links, the  $K$  sink nodes must transmit known pilot symbols over a period  $\tau_p \geq KN_r$ . Similar to the analysis done in [19], we let each sink independently and simultaneously transmit an  $N_r \times \tau_p$  matrix of pilots  $\Phi_i$  such that  $\Phi_i \Phi_k^* = \delta_{ik} \mathbf{I}_{N_r}$  shown to be optimal in [29]. This training phase only requires synchronization.

Let  $\hat{\mathbf{Y}}_i = [\hat{\mathbf{y}}_i[1] \dots \hat{\mathbf{y}}_i[\tau_p]]$  be the  $N_t \times \tau_p$  received training matrix at source node  $i$ . Then the received training is

$$\hat{\mathbf{Y}}_i = \sqrt{\frac{\tau_p P_f}{N_r}} \sum_{k=1}^K \mathbf{G}_{k,i} \Phi_k + \mathbf{V}_i, \quad \forall i$$

where  $\mathbf{V}_i$  is an  $N_t \times \tau_p$  matrix of i.i.d  $\mathcal{CN}(0, \sigma^2)$  noise elements. Using its received training, each source  $i$  locally computes MMSE estimates of its channels given by

$$\hat{\mathbf{G}}_{k,i} = \frac{\sqrt{\frac{\tau_p P_f}{N_r}}}{\sigma^2 + \frac{\tau_p P_f}{N_r}} \hat{\mathbf{Y}}_i \Phi_k^* \quad \forall k. \quad (5)$$

Since  $\hat{\mathbf{G}}_{k,i}$  are MMSE estimates of Gaussian variables corrupted by Gaussian noise, this results in  $\hat{\mathbf{G}}_{k,i} \sim \mathcal{CN}\left(0, \frac{\tau_p P_f}{\sigma^2 + \frac{\tau_p P_f}{N_r}}\right)$  and  $\tilde{\mathbf{G}}_{k,i} = \mathbf{G}_{k,i} - \hat{\mathbf{G}}_{k,i}$  with i.i.d  $\mathcal{CN}\left(0, \frac{\sigma^2}{\sigma^2 + \frac{\tau_p P_f}{N_r}}\right)$  elements.

2) *Analog CSI Feedback*: After reverse link training, each sink node  $i$  independently sends its unquantized uncoded estimates of  $\mathbf{H}_{i,k} \forall k$  over a period  $\tau_c$ . To have the sink nodes feedback their CSI simultaneously and orthogonally, each sink post multiplies its  $N_r \times KN_t$  feedback matrix  $[\mathbf{H}_{i,1} \dots \mathbf{H}_{i,K}]$  with a  $KN_t \times \tau_c$  matrix  $\Psi_i$  such that  $\Psi_i \Psi_i^* = \delta_{i,k} \mathbf{I}_{KN_t}$  [19]. This is a general orthogonal structure that can capture the case of orthogonality in time and requires  $\tau_c \geq K^2 N_t$ . The transmitted  $N_r \times \tau_c$  feedback matrix  $\overleftarrow{\mathbf{X}}_i$  from sink  $i$  can be written as

$$\overleftarrow{\mathbf{X}}_i = \sqrt{\frac{\tau_c P_f}{KN_t N_r}} [\mathbf{H}_{i,1} \dots \mathbf{H}_{i,K}] \Psi_i.$$

The concatenated  $KN_t \times \tau_c$  matrix of received feedback is then given by

$$\overleftarrow{\mathbf{Y}}_c = \sqrt{\frac{\tau_c P_f}{KN_t N_r}} \sum_{i=1}^K \begin{bmatrix} \mathbf{G}_{i,1} \\ \vdots \\ \mathbf{G}_{i,K} \end{bmatrix} [\mathbf{H}_{i,1} \dots \mathbf{H}_{i,K}] \Psi_i + \mathbf{V}$$

where  $\mathbf{V}$  is now a  $KN_t \times \tau_c$  noise matrix.

To estimate the forward channels  $\mathbf{H}_{i,k}$ , the source nodes first isolate the training from sink node  $i$  by post multiplying their received training by  $\Psi_i^*$  which gives

$$\overleftarrow{\mathbf{Y}}_c \Psi_i^* = \sqrt{\frac{\tau_c P_f}{KN_t N_r}} \underbrace{\begin{bmatrix} \mathbf{G}_{i,1} \\ \vdots \\ \mathbf{G}_{i,K} \end{bmatrix}}_{\mathbf{G}_i} \underbrace{[\mathbf{H}_{i,1} \dots \mathbf{H}_{i,K}]}_{\mathbf{H}_i} + \mathbf{V} \Psi_i^*. \quad (6)$$

To simplify the analysis in Section III-B, we assume that the complete received feedback matrix  $\overleftarrow{\mathbf{Y}}_c$  is shared among sources. Then, assuming  $KN_t \geq N_r$ , sources compute a common least squares estimate  $\hat{\mathbf{H}}_i$  of  $\mathbf{H}_i$  given by

$$\begin{aligned} \hat{\mathbf{H}}_i &= \sqrt{\frac{KN_t N_r}{\tau_c P_f}} \left( \hat{\mathbf{G}}_i^* \hat{\mathbf{G}}_i \right)^{-1} \hat{\mathbf{G}}_i^* \overleftarrow{\mathbf{Y}}_c \Psi_i^* \\ &= \underbrace{\mathbf{H}_i}_{\text{Real Channel}} + \underbrace{\tilde{\mathbf{H}}_i}_{\text{Error}}, \end{aligned}$$

where  $\hat{\mathbf{G}}_i$  is the MMSE estimate of  $\mathbf{G}_i$ . Such node cooperation is only realistic in certain cases



such as IA for cellular systems for example. At the end of this section, we provide alternative non-cooperative approaches that we show in Section V perform close to this special case.

The error in the estimates of  $\mathbf{H}_i$  can then be written as

$$\tilde{\mathbf{H}}_i = \left( \hat{\mathbf{G}}_i^* \hat{\mathbf{G}}_i \right)^{-1} \hat{\mathbf{G}}_i^* \left( \sqrt{\frac{K N_t N_r}{\tau_c P_f}} \mathbf{V} + \tilde{\mathbf{G}}_i \mathbf{H}_i \right).$$

which makes it clear that the error in the estimate consists of two error terms: the first due to noisy feedback and the second due to a noisy estimate of the feedback channel. To quantify the effect of the error on the achieved sum rate, we derive the variance of the error term introduced by analog feedback. Recall that the elements of  $\mathbf{H}_{i,k}$  are  $\mathcal{CN}(0, 1)$ , those of  $\mathbf{V}$  are  $\mathcal{CN}(0, \sigma^2)$ , and those of  $\tilde{\mathbf{G}}_i$  are  $\mathcal{CN}(0, \frac{\sigma^2}{\sigma^2 + \frac{\tau_p P_f}{N_r}})$ . As a result, the error term  $\tilde{\mathbf{G}}_i \mathbf{H}_i$  due to the reverse channel estimation has independent elements with a variance of  $\frac{N_r \sigma^2}{\sigma^2 + \frac{\tau_p P_f}{N_r}}$ . Similarly to [19] we see that the covariance of each column of  $\tilde{\mathbf{H}}_i$  denoted  $\tilde{\mathbf{H}}_i^{(\ell)}$ , conditioned on  $\hat{\mathbf{G}}_i$  is

$$\text{Cov}(\tilde{\mathbf{H}}_i^{(\ell)} | \hat{\mathbf{G}}_i) = \left( \frac{K N_t N_r \sigma^2}{\tau_c P_f} + \frac{N_r \sigma^2}{\sigma^2 + \frac{\tau_p P_f}{N_r}} \right) \left( \hat{\mathbf{G}}_i^* \hat{\mathbf{G}}_i \right)^{-1}.$$

Since the elements of the MMSE estimate  $\hat{\mathbf{G}}_i$  are Gaussian and uncorrelated, the diagonal elements of  $\left( \hat{\mathbf{G}}_i^* \hat{\mathbf{G}}_i \right)^{-1}$  are reciprocals of scaled chi-squared random variables with  $2(K N_t - N_r + 1)$  degrees of freedom. As a result, the mean square error,  $\sigma_f^2$ , in the elements of  $\hat{\mathbf{H}}_{i,k}$  is given by

$$\sigma_f^2 = \frac{\sigma^2}{(K N_t - N_r) P_f} \left( \frac{N_r^2}{\tau_p} + \frac{K N_t N_r}{\tau_c} \left( 1 + \frac{N_r \sigma^2}{\tau_p P_f} \right) \right). \quad (7)$$

At high SNR this gives

$$\sigma_f^2 \approx \frac{\sigma^2 \left( \frac{N_r^2}{\tau_p} + \frac{K N_t N_r}{\tau_c} \right)}{(K N_t - N_r) P_f}. \quad (8)$$

Having computed feedback error, we return to the assumption on node cooperation. Cooperation simplifies Section III-B by making a common channel estimate known to all users. Such cooperation may not always be possible. We present two alternative practical strategies, outline the changes they incur to the feedback process, and discuss their shortcomings:

- **Centralized processor:** In this scheme, reverse link training, estimation and feedback transmission remain unchanged. In this strategy however, the  $K N_t \times \tau_c$  matrix  $\overleftarrow{\mathbf{Y}}_c$  is not shared, and instead one of the nodes calculates an estimate of the feedback information based on

its locally observed  $N_t \times \tau_c$  matrix of feedback symbols (i.e. only its own rows of  $\overleftarrow{\mathbf{Y}}_c$ ). This can be done as long as  $N_t \geq N_r$ <sup>1</sup> and results in  $\sigma_f^2 \approx \frac{\sigma^2}{(N_t - N_r)P_f} \left( \frac{N_r^2}{\tau_p} + \frac{KN_t N_r}{\tau_c} \right)$ . This source then calculates the IA solution required and feeds it forward, again using “analog feedforward”, to all other sources [30].

- **Distributed processing:** In this strategy reverse link training, estimation and feedback transmission again remain unchanged. In this scheme, however, *each* source locally calculates its own estimate of all the channels being fed back, using the  $N_t \times \tau_c$  symbols it independently receives. Therefore, each node will obtain a different perturbed estimate of the forward channels with the same  $\sigma_f^2$  as in the centralized processor case. Each node then locally calculates a complete set of perturbed IA precoders based on its channel estimates.

We expand on the performance of these two alternative practical strategies in Section V.

### B. Multiplexing Gain with Analog Feedback

To characterize the performance of IA with analog feedback, we examine the rate loss [18] incurred by naive IA where channel estimates are used to calculate the columns of the precoders,  $\hat{\mathbf{f}}_i^m \forall i, m$  and combiners  $\hat{\mathbf{w}}_i^m \forall i, m$ . Therefore, such transmit and receive vectors satisfy

$$(\hat{\mathbf{w}}_i^m)^* \hat{\mathbf{H}}_{i,k} \hat{\mathbf{f}}_k^\ell = 0, \quad \forall (k, \ell) \neq (i, m) \quad (9)$$

$$\left| (\hat{\mathbf{w}}_i^m)^* \hat{\mathbf{H}}_{i,i} \hat{\mathbf{f}}_i^m \right| \geq c > 0, \quad \forall i, m, \quad (10)$$

and can be found for a feasible degree of freedom vector  $\mathbf{d} = [d_1, \dots, d_K]$ . As stated earlier, sinks need not use such per-stream combiners, and can employ a number of other receiver designs based on their channel knowledge. In Section V we show the performance of a joint decoder which whitens interference using its covariance [31]. Such receiver CSI can be acquired blindly [32] or via additional short training or silent phases. For the sake of the following rate loss analysis, we abstract the calculation of receiver CSI and assume that sinks have learned the combiners  $\hat{\mathbf{w}}_i^m$ .

The mean loss in sum-rate is then defined as  $\Delta R_{sum} \triangleq \mathbb{E}_{\mathbf{H}} R_{sum} - \mathbb{E}_{\mathbf{H}} \hat{R}_{sum}$ , where  $\mathbb{E}_{\mathbf{H}} R_{sum}$  is the average sum rate from IA with perfect CSI, with instantaneous rate given in (4), and

<sup>1</sup>The restriction  $N_t \geq N_r$  can be relaxed at the expense of a longer  $\tau_c$  by assigning each user several spreading matrices  $\Psi$ .

$\mathbb{E}_{\mathbf{H}} \widehat{R}_{sum}$  is the rate achieved with imperfect CSI and the vectors in (9), (10). Before proving the main result on rate loss, we will need the following lemma.

*Lemma 1:* The desired signal powers,  $|(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m|^2$  and  $|(\widehat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \widehat{\mathbf{f}}_i^m|^2$ , resulting from IA with perfect or imperfect CSI respectively are identically and exponentially distributed.

*Proof:* See Appendix A. ■

*Theorem 2:* The sum rate loss experienced by IA on the  $K$ -user  $N_r \times N_t$  interference channel with imperfect CSI obtained via the analog feedback strategy described in Section III-A is upper bounded by a constant when  $P_f$  scales with the transmit power  $P$ .

*Proof:* Let the  $K$ -user  $N_r \times N_t$  interference channel use the analog feedback scheme presented to achieve a vector of multiplexing gains  $\mathbf{d}$ . Using the imperfect vectors  $\widehat{\mathbf{f}}_i^m$  and  $\widehat{\mathbf{w}}_i^m$  respectively, the input-output relationship at the output of a linear zero-forcing receiver is

$$(\widehat{\mathbf{w}}_i^m)^* \mathbf{y}_i = (\widehat{\mathbf{w}}_i^m)^* \sqrt{\frac{P}{d_i}} \mathbf{H}_{i,i} \widehat{\mathbf{f}}_i^m s_i^m + \sum_{\ell \neq m} (\widehat{\mathbf{w}}_i^m)^* \sqrt{\frac{P}{d_i}} \mathbf{H}_{i,i} \widehat{\mathbf{f}}_i^\ell s_i^\ell + \sum_{k \neq i} \sum_{\ell=1}^{d_k} (\widehat{\mathbf{w}}_i^m)^* \sqrt{\frac{P}{d_k}} \mathbf{H}_{i,k} \widehat{\mathbf{f}}_k^\ell s_k^\ell + (\widehat{\mathbf{w}}_i^m)^* \mathbf{v}_i.$$

Using this received signal and the instantaneous rate expression in (4) gives the following upper bound on mean loss in sum rate:

$$\begin{aligned} \Delta R_{sum} &= \mathbb{E}_{\mathbf{H}} \sum_{i,m} \log_2 \left( 1 + \frac{\frac{P}{d_i} |(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m|^2}{\sigma^2} \right) - \mathbb{E}_{\mathbf{H}, \widehat{\mathbf{H}}} \sum_{i,m} \log_2 \left( 1 + \frac{\frac{P}{d_i} |(\widehat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \widehat{\mathbf{f}}_i^m|^2}{\mathcal{I}_{i,m} + \sigma^2} \right) \\ &= \mathbb{E}_{\mathbf{H}} \sum_{i,m} \log_2 \left( 1 + \frac{\frac{P}{d_i} |(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m|^2}{\sigma^2} \right) \\ &\quad - \mathbb{E}_{\mathbf{H}, \widehat{\mathbf{H}}} \sum_{i,m} \log_2 \left( 1 + \frac{\mathcal{I}_{i,m} + \frac{P}{d_i} |(\widehat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \widehat{\mathbf{f}}_i^m|^2}{\sigma^2} \right) \\ &\quad + \mathbb{E}_{\mathbf{H}, \widehat{\mathbf{H}}} \sum_{i,m} \log_2 \left( 1 + \frac{\mathcal{I}_{i,m}}{\sigma^2} \right) \\ &\stackrel{(a)}{\leq} \mathbb{E}_{\mathbf{H}, \widehat{\mathbf{H}}} \sum_{i,m} \log_2 \left( 1 + \frac{\mathcal{I}_{i,m}}{\sigma^2} \right) \stackrel{(b)}{\leq} \sum_{i,m} \log_2 \left( 1 + \frac{\mathbb{E}_{\mathbf{H}, \widehat{\mathbf{H}}} \mathcal{I}_{i,m}}{\sigma^2} \right). \end{aligned}$$

where (a) follows from Lemma 1 which along with  $\mathcal{I}_{i,m} \geq 0$  implies that  $\frac{P}{d_i} |(\widehat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \widehat{\mathbf{f}}_i^m|^2 + \mathcal{I}_{i,m}$  stochastically dominates  $\frac{P}{d_i} |(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m|^2$  and (b) follows from Jensen's inequality [18].

Since (9) and (10) are satisfied, however, the total interference term  $\mathcal{I}_{i,m}$  can be simplified to

include only residual interference due to the channel estimation errors  $\tilde{\mathbf{H}}_{i,\ell}$ . Therefore,  $\Delta R_{sum}$  can be further upper bounded by noticing that

$$\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \left| (\hat{\mathbf{w}}_i^m)^* (\hat{\mathbf{H}}_{i,k} - \tilde{\mathbf{H}}_{i,k}) \hat{\mathbf{f}}_k^\ell \right|^2 = \mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \left| (\hat{\mathbf{w}}_i^m)^* (\tilde{\mathbf{H}}_{i,k}) \hat{\mathbf{f}}_k^\ell \right|^2 \leq \mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \|\tilde{\mathbf{H}}_{i,k}\|_F^2, \quad \forall k, \forall \ell \neq m \quad (11)$$

which gives

$$\Delta R_{sum} \leq \sum_{i,m} \log_2 \left( 1 + \frac{1}{\sigma^2} \sum_{\ell=1}^K \frac{P}{d_\ell} (d_\ell - \delta_{i,\ell}) \mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \|\tilde{\mathbf{H}}_{i,\ell}\|_F^2 \right). \quad (12)$$

From (7), however, we have  $\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \|\tilde{\mathbf{H}}_{i,\ell}\|_F^2 = N_t N_r \sigma_f^2 = \frac{c(\tau_p, \tau_c) \sigma^2}{P_f}$ , where  $c(\tau_p, \tau_c)$  is a constant, independent of  $P_f$  at high enough SNR, given by

$$c(\tau_p, \tau_c) = N_t N_r \frac{\left( \frac{N_r^2}{\tau_p} + \frac{K N_t N_r (1+\epsilon)}{\tau_c} \right)}{(K N_t - N_r)}. \quad (13)$$

Combining (12), (13) and letting  $P_f = \alpha^{-1} P$  gives the final upper bound on throughput loss

$$\Delta R_{sum}(\tau_p, \tau_c) \leq \sum_{i,m} \log_2 \left( 1 + \frac{c(\tau_p, \tau_c) \sigma^2 P}{\sigma^2 P_f} \sum_{\ell=1}^K \frac{d_\ell - \delta_{i,\ell}}{d_\ell} \right) \leq \sum_i d_i \log_2 \left( 1 + \alpha c(\tau_p, \tau_c) \left( K - \frac{1}{d_i} \right) \right). \quad (14)$$

The bound has been presented at high SNR for simplicity of exposition only; (13) can be adapted for any SNR  $> 0$  by using (7) instead of (8).  $\blacksquare$

*Corollary 3:* IA with imperfect CSI obtained via the analog feedback strategy in Section III-A achieves the same average multiplexing gain as IA with perfect CSI.

*Proof:* This follows immediately from the constant loss in sum rate shown in Theorem 2 which becomes negligible as SNR is taken to infinity.  $\blacksquare$

In summary, Theorem 2 states that if the SNR on the reverse and forward link are comparable, the cost of imperfect CSI is a constant. This constant is a decreasing function of  $\tau_c$  and  $\tau_p$  and thus, we have written  $\Delta R_{sum}(\tau_p, \tau_c)$  in (14) to highlight this dependence. Note that while bounding leakage interference as  $\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \left| (\hat{\mathbf{w}}_i^m)^* (\tilde{\mathbf{H}}_{i,k}) \hat{\mathbf{f}}_k^\ell \right|^2 \leq \mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \|\tilde{\mathbf{H}}_{i,k}\|_F^2 = N_t N_r \sigma_f^2$  suffices to establish a constant rate loss, it is very conservative, and increasingly loose for larger systems. If we assume that  $\tilde{\mathbf{H}}_{i,k}$  has a bi-unitary invariant distribution, e.g. Gaussian, then by the same reasoning as in Appendix A we get  $\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \left| (\hat{\mathbf{w}}_i^m)^* (\tilde{\mathbf{H}}_{i,k}) \hat{\mathbf{f}}_k^\ell \right|^2 = \sigma_f^2$  and  $c(\tau_p, \tau_c)$  is replaced with  $c_2(\tau_p, \tau_c) = \frac{1}{K N_t - N_r} \left( \frac{N_r^2}{\tau_p} + \frac{K N_t N_r (1+\epsilon)}{\tau_c} \right)$ . Since SNRs are likely to be comparable in practice, analog feedback allows systems to overcome the problem of fast scaling complexity in digital feedback, which remains even if the actual feedback stage operates error-free close to capacity [18], [21]. In

case such SNR symmetry is not possible, however, systems become interference limited and multiplexing gain is reduced.

*Theorem 4:* IA on a  $K$ -user  $N_r \times N_t$  interference channel using analog feedback with  $P_f = \alpha P^\beta$  such that  $0 \leq \beta \leq 1$  achieves at least a  $\beta$ -fraction of its original multiplexing gain.

*Proof:* The multiplexing gain achieved by a system using naive IA and analog feedback with  $P_f = \alpha P^\beta$  can be written as

$$\begin{aligned}
M &= \lim_{P \rightarrow \infty} \frac{\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}}[\hat{R}_{sum}]}{\log_2(P)} = \lim_{P \rightarrow \infty} \frac{\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \sum_{i,m} \log_2 \left( 1 + \frac{\frac{P}{d_i} |(\hat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \hat{\mathbf{f}}_i^m|^2}{\mathcal{I}_{i,m} + \sigma^2} \right)}{\log_2(P)} \\
&= \lim_{P \rightarrow \infty} \frac{\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \sum_{i,m} \log_2 \left( \frac{P}{d_i} |(\hat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \hat{\mathbf{f}}_i^m|^2 + \mathcal{I}_{i,m} + \sigma^2 \right) - \mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \sum_{i,m} \log_2 (\mathcal{I}_{i,m} + \sigma^2)}{\log_2(P)} \\
&\stackrel{(a)}{\geq} \frac{\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \sum_{i,m} \log_2 \left( \frac{P}{d_i} |(\hat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \hat{\mathbf{f}}_i^m|^2 \right) - \sum_{i,m} \log_2 (\mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \mathcal{I}_{i,m} + \sigma^2)}{\log_2(P)} \\
&\stackrel{(b)}{\geq} \sum_{k=1}^K d_k - \lim_{P \rightarrow \infty} \frac{\sum_{i,m} \log_2 \left( \frac{(d_i-1)P}{d_i} \mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \|\tilde{\mathbf{H}}_{i,i}\|_F^2 + (K-1)P \mathbb{E}_{\mathbf{H}, \hat{\mathbf{H}}} \|\tilde{\mathbf{H}}_{i,k}\|_F^2 + \sigma^2 \right)}{\log_2(P)} \\
&\stackrel{(c)}{=} \sum_{k=1}^K d_k - \lim_{P \rightarrow \infty} \frac{\sum_{i,m} \log_2 \left( \frac{(d_i-1)P^{1-\beta} c(\tau_p, \tau_c) \sigma^2}{\alpha d_i} + \frac{(K-1)P^{1-\beta} c(\tau_p, \tau_c) \sigma^2}{\alpha} + \sigma^2 \right)}{\log_2(P)} = \beta \left( \sum_{k=1}^K d_k \right),
\end{aligned}$$

where (a) follows from disregarding interference in the first term and applying Jensen's to the second; (b) follows from (20) and the fact that each term in the first summation has a multiplexing gain of 1. Finally, (c) is due to the scaling of residual interference with  $P^{1-\beta}$ . ■

In the case where  $\beta = 0$ , feedback power is constant, and the interference limited system achieves zero multiplexing gain. In fact, the system simulations in Section V indicate that the sum-rate achieved is upper bounded by a constant. As for the case of  $\beta = 1$ , Theorem 4 confirms the preservation of full multiplexing gain proven in Corollary 3, but does not establish the  $O(1)$  loss in sum rate of Theorem 2. As for any  $0 < \beta < 1$ , Theorem 4 shows that analog feedback and interference alignment can still achieve linear sum rate scaling even when feedback power is much smaller than transmit power. This may be the case in certain non-homogeneous networks such as cellular networks in which mobile and base station powers do not match.

#### IV. DEGREES OF FREEDOM WITH OVERHEAD

While the analysis done in Section III-B is a good indicator of the cost of imperfect CSI obtained through training and analog feedback, it does not directly predict the expected throughput achieved by this strategy. Namely, the analysis done thus far neglects the cost of training and feedback overhead. In this section we define the expected throughput with overhead and use it to optimize training and feedback.

##### A. Definition of Overhead

As shown in Section III-B, the performance of IA is tightly related to the mean square error in the channel estimates at the transmitter. When operating in a time varying channel, training and feedback must be done periodically to ensure the validity of the channel estimate at the transmitter. Depending on the channel's coherence time, the overhead due to training and feedback may consume an arbitrarily large fraction of resources such as time or frequency slots, resulting in low net throughput. In this section we consider the case in which training, feedback, and data transmission are all orthogonal in time, in the same coherence time or frame  $T$  [15], [33]. Using this model for training and feedback overhead we compute the expected effective sum rate as

$$R_{eff}(\tau_p, \tau_c) = \left( \frac{T - (\tau_c + \tau_p)}{T} \right) (\mathbb{E}_{\mathbf{H}} R_{sum} - \Delta R_{sum}). \quad (15)$$

Using the rate loss expression in (14), we write the expected sum rate with overhead as

$$R_{eff}(\tau_p, \tau_c) \approx \left( \frac{T - (\tau_c + \tau_p)}{T} \right) \left( \mathbb{E}_{\mathbf{H}} R_{sum} - \sum_i d_i \log_2 \left( 1 + \alpha c_2(\tau_p, \tau_c) \left( K - \frac{1}{d_i} \right) \right) \right), \quad (16)$$

where we have used  $c_2(\tau_p, \tau_c)$  rather than the looser  $c(\tau_p, \tau_c)$ . As expected, and as can be seen from (15), insufficient training and feedback may result in poor channel estimates at the receiver, and thus a large loss in sum rate, whereas excessive training and feedback becomes too costly as a large portion of the frame is spent on overhead.

### B. Training and Feedback Optimization

Given the expression for system throughput with overhead, we propose solving the following optimization problem

$$\max_{\tau_p, \tau_c} \left( \frac{T - (\tau_c + \tau_p)}{T} \right) \left( \mathbb{E}_{\mathbf{H}} R_{sum} - \sum_i d_i \log_2 \left( 1 + \alpha \frac{\left( \frac{N_r^2}{\tau_p} + \frac{K N_t N_r}{\tau_c} \right)}{(K N_t - N_r)} \left( K - \frac{1}{d_i} \right) \right) \right), \quad (17)$$

over the set of feasible training and feedback lengths. The sum rate expression defined in (15) is not convex as it is defined on a non-convex non-continuous closed set of bounded integers, as would be the case if one were optimizing over the number of feedback bits. Nevertheless, we seek to maximize a continuous relaxation of the defined cost function by typical methods of convex optimization [21], [33].

The optimization can be simplified by realizing that for a fixed overhead length  $T_{total}$ , optimizing training and feedback lengths simplifies to minimizing  $c_2(\tau_p, \tau_c)$  and then finding the optimal overhead length. So given a fixed amount of overhead the solution of the first optimization step can be shown to be

$$\tau_p = \frac{N_r}{N_r + \sqrt{K N_t N_r}} T_{total}, \quad \tau_c = \frac{\sqrt{K N_t N_r}}{N_r + \sqrt{K N_t N_r}} T_{total}, \quad (18)$$

which gives the optimal value  $c_{opt} = \frac{(N_r + \sqrt{K N_t N_r})^2}{(K N_t - N_r) T_{total}}$ . Replacing  $c_{opt}$  into (17), the optimal training length  $T_{total}$  can now be found by solving

$$\begin{aligned} \frac{\delta}{\delta T_{total}} \left( \frac{T - T_{total}}{T} \right) \left( \mathbb{E}_{\mathbf{H}} R_{sum} - \sum_i d_i \log_2 \left( 1 + \frac{\alpha (K - 1/d_i) (N_r + \sqrt{K N_t N_r})^2}{(K N_t - N_r) T_{total}} \right) \right) &= 0 \\ -\frac{1}{T} \left( \mathbb{E}_{\mathbf{H}} R_{sum} - \sum_i d_i \log \left( 1 + \frac{\gamma_i}{T_{total}} \right) \right) + \frac{T - T_{total}}{T} \left( \sum \frac{d_i}{T_{total} (T_{total} + \gamma_i)} \right) &= 0 \end{aligned} \quad (19)$$

where  $\gamma_i = \frac{\alpha (K - 1/d_i) (N_r + \sqrt{K N_t N_r})^2}{(K N_t - N_r)}$ . Though the exact solution to this optimization cannot be found in closed form, it can be shown that  $T_{total}$  increases with  $\sqrt{T}$ , and *initially* decreases with the ratio transmit to feedback power  $\frac{P}{P_f}$ . Fig. 2 plots the optimal feedback vs. frame length, where we have solved the optimization in (19) numerically for a three  $2 \times 2$  user network with  $d_i = 1$ . Fig. 2 verifies the claimed scaling and shows that for a range of reasonable frame lengths, the solution to the optimization problem is less than the minimum length required to satisfy the dimensionality constraints on the training and feedback matrices, i.e. the optimal overhead is minimal for realistic frame sizes.

While it may not be surprising that longer frames can support more training [33], it's interesting to note that small mismatches in forward and feedback SNR initially decrease training and feedback lengths. Hence, a slightly noisy feedback channel does not require extra training to compensate. This, however, is not true of significantly poor feedback channels, as the optimal overhead length does in fact increase to improve the quality of CSI. It can also be shown that, all else fixed, the optimal training length decreases with the achieved sum rate or effectively SNR, making analog feedback especially efficient at high SNR. This is shown in Fig. 3.

## V. SIMULATION RESULTS

In this section we present simulation results that validate the claims and proofs given in Sections III-B and IV. We verify the results shown in Theorems 2 and 4 which state that as long as the transmit power on the feedback channel scales sufficiently with the power on the forward channel, the multiplexing gain achieved by perfect IA is preserved. To better show the performance of both IA and IA with analog feedback we remove the restriction of per-stream receivers and thus calculate the sum rate of a joint decoder which also treats interference as colored Gaussian noise

$$R_{sum} = \sum_{i=1}^K \log_2 \left| \mathbf{I} + \frac{P}{d_i} \mathbf{H}_{i,i} \mathbf{F}_i \mathbf{F}_i^* \mathbf{H}_{i,i}^* \left( \sigma^2 \mathbf{I} + \sum_{k \neq i} \frac{P}{d_k} \mathbf{H}_{i,k} \mathbf{F}_k \mathbf{F}_k^* \mathbf{H}_{i,k}^* \right)^{-1} \right|,$$

and the precoders are calculated given ideal or estimated CSI [31].

Fig. 4 shows the sum rate achieved by the IA algorithm in [3] in a 3 user  $5 \times 4$  system with  $d_i = 2 \forall i$ . We show performance with perfect CSI, scaling quality CSI where  $P_f = P/2$ , slower scaling CSI with  $P_f = P^\beta$  and  $\beta = 0.5$  and fixed quality CSI where the SNR on the feedback channel is fixed at 5dB. In all cases, minimum training and feedback lengths are used, i.e.  $\tau_p = 12$  and  $\tau_c = 45$ . Fig. 4 confirms that both perfect and scaling feedback exhibit the same sum rate scaling or degrees of freedom. This establishes the multiplexing gain optimality of using analog feedback. Fig. 4 also confirms the fact that the mean loss in sum rate at high enough SNR is indeed a constant independent of the forward channel SNR. Moreover, Fig. 4 shows that while the rate loss bound with  $c(\tau_p, \tau_c)$  suffices to establish a constant rate loss and multiplexing gain optimality, it is loose for larger systems as mentioned in Section III-B. Using  $c_2(\tau_p, \tau_c)$ , by assuming a biunitary invariant error distribution, gives a better characterization of



achieved sum rate. Moreover, both rate loss characterizations, as well as simulated rate loss, naturally decrease with  $\tau_p$  and  $\tau_c$  as stated in Section III-B; the trend can be seen in Fig. 5 for a  $2 \times 2$  system with  $P_f = P/10$ . When perfect scaling feedback is not possible, Fig. 4 shows the slower yet linear scaling in the case of  $P_f = P^\beta$  which verifies the result shown in Theorem 4 on the preservation of a fraction of the system's original multiplexing gain. Finally, for the case of fixed feedback quality, multiplexing gain is zero and the sum rate saturates at high SNR.

Fig. 4 also shows the performance of the distributed processing approach introduced in Section III-A. As stated earlier, the cooperation assumed in Section III-A is only practical in certain cases such as cellular systems. If cooperation is not possible, either a central node can calculate the IA solution and feed it forward to the other sources, or nodes calculate precoders independently. The analysis of the centralized processor strategy is straightforward. It only adds Gaussian noise to the precoders due to feed forward; i.e. all nodes will use noisy versions of the same vector,  $\hat{\mathbf{f}}_i^\ell + \tilde{\mathbf{f}}_i^\ell$  and  $\hat{\mathbf{w}}_i^\ell + \tilde{\mathbf{w}}_i^\ell$ . So, the interference terms are now given by

$$(\hat{\mathbf{w}}_i^\ell + \tilde{\mathbf{w}}_i^\ell)^* \underbrace{(\hat{\mathbf{H}}_{i,k} + \tilde{\mathbf{H}}_{i,k})}_{\mathbf{H}_{i,k}} (\hat{\mathbf{f}}_i^\ell + \tilde{\mathbf{f}}_i^\ell) = (\hat{\mathbf{w}}_i^\ell)^* \underbrace{\tilde{\mathbf{H}}_{i,k}}_{P_f^{-1}} \hat{\mathbf{f}}_k^\ell + (\hat{\mathbf{w}}_i^\ell)^* \mathbf{H}_{i,k} \underbrace{\tilde{\mathbf{f}}_i^\ell}_{P_f^{-1}} + \underbrace{(\tilde{\mathbf{w}}_i^\ell)^*}_{P_f^{-1}} \mathbf{H}_{i,k} \hat{\mathbf{f}}_i^\ell + \underbrace{(\tilde{\mathbf{w}}_i^\ell)^*}_{P_f^{-1}} \mathbf{H}_{i,k} \underbrace{\tilde{\mathbf{f}}_i^\ell}_{P_f^{-1}}.$$

All errors in this case again decay with SNR as indicated above, and similar bounds on rate loss can be readily derived. In the case of distributed processing, the extra loss is due to the mismatch of CSI between nodes; IA solutions based on different perturbations of the same channels will be mismatched leading to extra interference leakage. Fig. 4, however, shows that the extra loss due to distributed processing is small, and no degrees of freedom are lost. The performance of distributed processing is not theoretically surprising. Since IA solutions and algorithms depend heavily on invariant and singular subspaces, Wedin's theorem [34], [35] and results in [36] can be used to tractably bound the angles between singular and invariant subspaces of perturbed matrices. Using the bounds in [35] we can see that even if nodes compute precoders using different perturbed channels, the angle (or error) between the different precoders calculated is bounded by a linear function of  $\|\tilde{\mathbf{H}}_{i,k}\|_F^2$ . Therefore, errors still decay with feedback power. Such decay is all that is needed to prove the multiplexing gain preservation and a constant rate loss. Computing a tight constant bound on  $\Delta R_{sum}$ , however, becomes significantly more involved. The small loss due to distributed processing, and the fact that it does not need extra overhead, make it a practical and viable approach.

Finally, we simulate the system's total throughput according to the overhead model presented

in Section IV. Fig. 6 shows that when the forward and reverse channel SNR scale together, the optimal feedback length is close to the theoretical minimum required for a frame length of 2,000, as predicted by Fig. 2 even for the relatively poor feedback channel ( $P_f = P/100$ ) considered. Although analog feedback provides poor overhead scaling with frame length, the effect of this scaling is little since overhead is minimal for practical frames [33]. Finally, Fig 6 shows the optimal feedback length assuming fixed training for ease of exposition. We know, however, from Section IV, that this is not optimal. Fig. 7 shows that while generally less resources are used for training, both training and feedback must scale together for maximum throughput.

## VI. CONCLUSIONS

In this paper we proposed a low overhead feedback strategy for the interference channel. We showed that when combined with interference alignment, analog feedback can achieve full multiplexing gain when the forward and reverse channel SNR levels are comparable. When such symmetry is not possible, we showed a fraction of the degrees of freedom is retained. Thus a main benefit of analog feedback is that the cost of imperfect channel knowledge at the transmitter is bounded and quickly becomes negligible at high SNR. The mild requirement of comparable feedback and transmit power implies that analog feedback performs well with constant overhead, in the high SNR regime where IA is optimal. In addition to quantifying the cost of imperfect CSI, we show the scaling of required overhead with several network variables such as SNR and frame length. In simulation, we show that the throughput loss due to the overhead of training and analog feedback is often minimal.

## APPENDIX A

### PROOF OF LEMMA 1

Consider [1], [3], [6] or similar solutions to finding unitary IA precoders  $\mathbf{F}_i = [\mathbf{f}_i^1, \dots, \mathbf{f}_i^{d_i}]$ . In all such solutions,  $\mathbf{F}_i \forall i$  are functions of all interfering channels  $\mathbf{H}_{k,\ell} \forall k, \ell, k \neq \ell$  only. Since  $\mathbf{F}_i$  is not a function of  $\mathbf{H}_{i,i}$ , and channels are independent across users, this implies that  $\mathbf{F}_i$  is independent of  $\mathbf{H}_{i,i} \forall i$ . Therefore,  $\mathbf{H}_{i,i}\mathbf{f}_i^m$  are Gaussian vectors with covariance  $\mathbb{E}[\mathbf{H}_{i,i}\mathbf{f}_i^m(\mathbf{f}_i^m)^*\mathbf{H}_{i,i}^*] = \text{trace}(\mathbf{f}_i^m(\mathbf{f}_i^m)^*)\mathbf{I}_{N_r} = \mathbf{I}_{N_r}$  [37, Ch. 21]. Similarly, due to the unitary precoders,  $\mathbb{E}[\mathbf{H}_{i,i}\mathbf{f}_i^m(\mathbf{f}_i^\ell)^*\mathbf{H}_{i,i}^*] = \text{trace}(\mathbf{f}_i^m(\mathbf{f}_i^\ell)^*)\mathbf{I}_{N_r} = \mathbf{0}_{N_r}$ , and therefore  $\mathbf{H}_{i,i}\mathbf{f}_i^m$  and  $\mathbf{H}_{i,i}\mathbf{f}_i^\ell$  are independent  $\forall \ell \neq m$ . Vectors  $\hat{\mathbf{f}}_i^m$  are calculated similarly based on  $\hat{\mathbf{H}}_{k,\ell} \forall k, \ell, k \neq \ell$  and are thus independent of both  $\hat{\mathbf{H}}_{i,i}$  and  $\mathbf{H}_{i,i}$ , thus  $\mathbf{H}_{i,i}\hat{\mathbf{f}}_i^\ell$  satisfy the same properties as  $\mathbf{H}_{i,i}\mathbf{f}_i^\ell$ .

The combiners  $\mathbf{w}_i^m$  must now satisfy (2); for continuously distributed i.i.d channels (3) will be satisfied automatically [1]. This can be done by letting

$$\mathbf{w}_i^m = \mathbf{U}_{\min} \left( \left[ \mathbf{H}_{i,1} \mathbf{F}_1, \dots, \mathbf{H}_{i,i-1} \mathbf{F}_{i-1}, \mathbf{H}_{i,i} \mathbf{F}_i^{(m)}, \mathbf{H}_{i,i+1} \mathbf{F}_{i+1}, \dots, \mathbf{H}_{i,K} \mathbf{F}_K \right] \right), \quad (20)$$

where  $\mathbf{F}_i^{(m)} = [\mathbf{f}_1^1, \dots, \mathbf{f}_1^{m-1}, \mathbf{f}_1^{m+1}, \dots, \mathbf{f}_1^{d_i}]$ , and  $\mathbf{U}_{\min}(\mathbf{A})$  extracts the least dominant left singular vector of  $\mathbf{A}$ . Given that IA is feasible, the interference matrix in (20) will span at most  $N_r - 1$  dimensions (at most  $N_r - d_i$  and  $d_i - 1$  dimensions of inter-user and inter-stream interference respectively). Hence,  $\mathbf{w}_i^m$  will always correspond to a 0 singular value, thus satisfying (2).

So  $\mathbf{w}_i^m$  are a function of vectors  $\mathbf{H}_{i,k} \mathbf{f}_k^\ell \forall k \neq i, \forall \ell$  and  $\mathbf{H}_{i,i} \mathbf{f}_i^\ell, \forall \ell \neq m$ . But  $\mathbf{H}_{i,k} \mathbf{f}_k^\ell \forall k \neq i, \forall \ell$  are independent of  $\mathbf{H}_{i,i} \mathbf{f}_i^m$  due to the independence of channels, and  $\mathbf{H}_{i,i} \mathbf{f}_i^\ell, \forall \ell \neq m$  are independent of  $\mathbf{H}_{i,i} \mathbf{f}_i^m$  as shown earlier. Thus  $\mathbf{w}_i^m$  is independent of  $\mathbf{H}_{i,i} \mathbf{f}_i^m$ . Since  $\mathbf{w}_i^m$  and  $\mathbf{H}_{i,i} \mathbf{f}_i^m$  are independent, and since  $\mathbf{H}_{i,i} \mathbf{f}_i^m$  has a unitary invariant Gaussian distribution as shown earlier [38], we can perform a change of basis such that  $\mathbf{w}_i^m = [1, 0, \dots, 0]^*$ . After this change of basis  $(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m$  is simply the first element of  $\mathbf{H}_{i,i} \mathbf{f}_i^m$  which we have shown above is  $\mathcal{CN}(0, 1)$ . Therefore, the terms  $|(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m|^2$  are exponentially distributed with parameter 1.

The vectors  $\hat{\mathbf{w}}_i^m$  are again given by (20) using  $\hat{\mathbf{H}}_{k,\ell}$  and  $\hat{\mathbf{f}}_k^\ell$  instead of  $\mathbf{H}_{k,\ell}$  and  $\mathbf{f}_k^\ell$  respectively, and are thus again functions of variables independent of both  $\hat{\mathbf{H}}_{i,i} \hat{\mathbf{f}}_i^m$  and  $\mathbf{H}_{i,i} \hat{\mathbf{f}}_i^m$ . Thus  $\hat{\mathbf{w}}_i^m$  are independent of  $\mathbf{H}_{i,i} \hat{\mathbf{f}}_i^m$  and again a change of basis reveals that  $(\hat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \hat{\mathbf{f}}_i^m$  are  $\mathcal{CN}(0, 1)$  and  $|(\hat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \hat{\mathbf{f}}_i^m|^2$  are exponentially distributed with parameter 1. Therefore,  $|(\mathbf{w}_i^m)^* \mathbf{H}_{i,i} \mathbf{f}_i^m|^2$  and  $|(\hat{\mathbf{w}}_i^m)^* \mathbf{H}_{i,i} \hat{\mathbf{f}}_i^m|^2$  are identically and exponentially distributed with parameter 1.

## REFERENCES

- [1] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [2] T. Gou and S. Jafar, "Degrees of freedom of the K user MxN MIMO interference channel," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6040–6057, Dec. 2010.
- [3] S. Peters and R. W. Heath, Jr., "Cooperative algorithms for MIMO interference channels," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 1, pp. 206–218, Jan. 2011.
- [4] C. Yetis, T. Gou, S. Jafar, and A. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4771–4782, Sept. 2010.
- [5] O. El Ayach, S. Peters, and R. W. Heath, Jr., "The feasibility of interference alignment over measured MIMO-OFDM channels," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 9, pp. 4309–4321, Nov. 2010.
- [6] K. Gomadam, V. Cadambe, and S. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," *Proc. of IEEE Global Telecommunications Conference*, pp. 1–6, Nov. 30 - Dec. 4 2008.
- [7] C. Shi, R. A. Berry, and M. L. Honig, "Adaptive Beamforming in Interference Networks via Bi-Directional Training," *ArXiv preprint arXiv:1003.4764*, Mar. 2010. [Online]. Available: <http://arxiv.org/abs/1003.4764>
- [8] D. Schmidt, C. Shi, R. Berry, M. Honig, and W. Utschick, "Minimum mean squared error interference alignment," *Proc. of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pp. 1106–1110, Nov. 2009.

- [9] S. W. Choi, S. Jafar, and S.-Y. Chung, "On the beamforming design for efficient interference alignment," *IEEE Communications Letters*, vol. 13, no. 11, pp. 847–849, Nov. 2009.
- [10] R. Tresch, M. Guillaud, and E. Riegler, "On the achievability of interference alignment in the K-user constant MIMO interference channel," *Proc. IEEE/SP 15th Workshop on Statistical Signal Processing*, pp. 277–280, Aug. 31–Sept. 3 2009.
- [11] I. Santamaría, O. González, R. W. Heath, Jr., and S. W. Peters, "Maximum sum-rate interference alignment algorithms for MIMO channels," *Proc. of IEEE Global Telecommunications Conference*, Dec. 2010.
- [12] D. S. Papailiopoulos and A. G. Dimakis, "Interference alignment as a rank constrained rank minimization," *Arxiv preprint arXiv:1010.0476*, 2010. [Online]. Available: <http://arxiv.org/abs/1010.0476>
- [13] V. Jungnickel, U. Kruger, G. Istoc, T. Haustein, and C. von Helmolt, "A MIMO system with reciprocal transceivers for the time-division duplex mode," *Proc. of IEEE Antennas and Propagation Society International Symposium*, vol. 2, pp. 1267–1270, Jun. 2004.
- [14] M. Guillaud, D. Slock, and R. Knopp, "A practical method for wireless channel reciprocity exploitation through relative calibration," *Proc. of the Eighth International Symposium on Signal Processing and Its Applications*, pp. 403–406, Aug. 2005.
- [15] S. Peters and R. W. Heath, Jr., "Orthogonalization to reduce overhead in MIMO interference channels," *Proc. of International Zurich Seminar*, pp. 126–129, Mar. 2010.
- [16] H. Bolcskei and J. Thukral, "Interference alignment with limited feedback," *Proc. IEEE International Symposium on Information Theory*, pp. 1759–1763, Jun. 28–Jul. 3 2009.
- [17] R. Krishnamachari and M. Varanasi, "Interference Alignment Under Limited Feedback for MIMO Interference Channels," *Arxiv preprint arXiv:0911.5509*, 2009. [Online]. Available: <http://arxiv.org/abs/0911.5509>
- [18] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.
- [19] T. Marzetta and B. Hochwald, "Fast transfer of channel state information in wireless systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1268–1278, Apr. 2006.
- [20] D. Samardzija and N. Mandayam, "Unquantized and uncoded channel state information feedback in multiple-antenna multiuser systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1335–1345, Jul. 2006.
- [21] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.
- [22] M. M. Shانهchi, R. Porat, and U. Erez, "Comparison of practical feedback algorithms for multiuser MIMO," *IEEE Transactions on Communications*, vol. 58, no. 8, pp. 2436–2446, Aug. 2010.
- [23] O. El Ayach and R. W. Heath, Jr., "Interference alignment with analog CSI feedback," *Proc. of IEEE Military Communications Conference*, pp. 1644–1648, Oct. 31–Nov. 3 2010.
- [24] E. Martinian, "Waterfilling gains  $O(1/\text{SNR})$  at high SNR," *Unpublished notes available from <http://www.csua.berkeley.edu/~emin/research/wfill.pdf>*.
- [25] C. Rose, S. Ulukus, and R. Yates, "Wireless systems and interference avoidance," *IEEE Transactions on Wireless Communications*, vol. 1, no. 3, pp. 415–428, Jul. 2002.
- [26] D. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.
- [27] B. Nosrat-Makouei, J. Andrews, and R. W. Heath, Jr., "A simple SINR characterization for linear interference alignment over uncertain MIMO channels," *Proc. of IEEE Int. Symposium on Inf. Theory*, pp. 2288–2292, Jun. 2010.
- [28] G. Caire, N. Jindal, and S. Shamai, "On the required accuracy of transmitter channel state information in multiple antenna broadcast channels," *Proc. of the Forty-First Asilomar Conference on Signals, Systems and Computers*, pp. 287–291, Nov. 2007.
- [29] T. Marzetta, "BLAST Training: Estimating Channel Characteristics for High-Capacity Space-Time Wireless," *Proc. of 37th Annual Allerton Conference on Communications, Control, and Computing*, Sept. 1999.
- [30] C. Chae, D. Mazzarese, T. Inoue, and R. W. Heath, Jr., "Coordinated beamforming for the multiuser MIMO broadcast channel with limited feedforward," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 6044–6056, Dec. 2008.
- [31] R. Blum, "MIMO capacity with interference," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 793–801, Jun. 2003.
- [32] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Transactions on Information Theory*, vol. 41, no. 4, pp. 944–960, Jul. 1995.
- [33] M. Kobayashi, N. Jindal, and G. Caire, "Training and Feedback Optimization for Multiuser MIMO Downlink," *to appear in IEEE Transactions on Communications*, 2011. [Online]. Available: <http://arxiv.org/abs/0912.1987>
- [34] P. Wedin, "Perturbation bounds in connection with singular value decomposition," *BIT Numerical Mathematics*, vol. 12, no. 1, pp. 99–111, 1972.
- [35] G. Stewart, "Perturbation theory for the singular value decomposition," *SVD and Signal Processing II, Algorithms, Analysis and Applications (Elsevier Science)*, pp. 99–109, 1991.
- [36] —, "Stochastic perturbation theory," *SIAM review*, vol. 32, no. 4, pp. 579–610, 1990.
- [37] A. Mathai and S. Provost, *Quadratic forms in random variables: theory and applications*. M. Dekker, New York, 1992.
- [38] A. Tulino and S. Verdú, *Random matrix theory and wireless communications*. Now Publishers Inc, 2004.

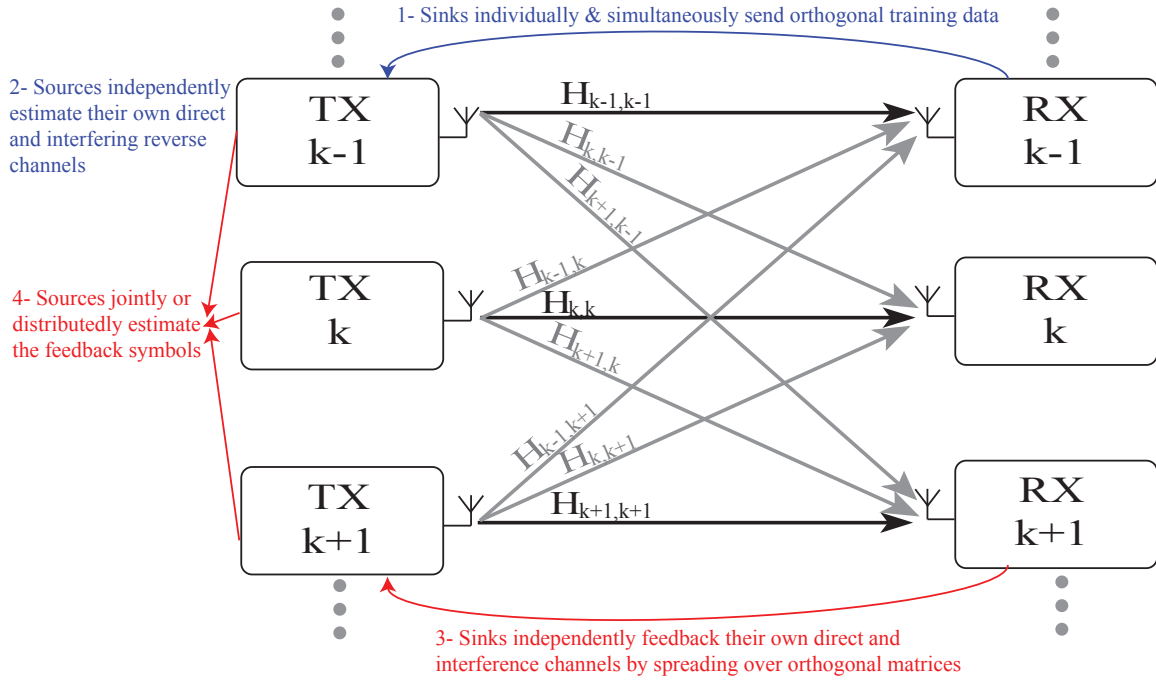


Fig. 1. K-User MIMO interference channel with analog feedback.

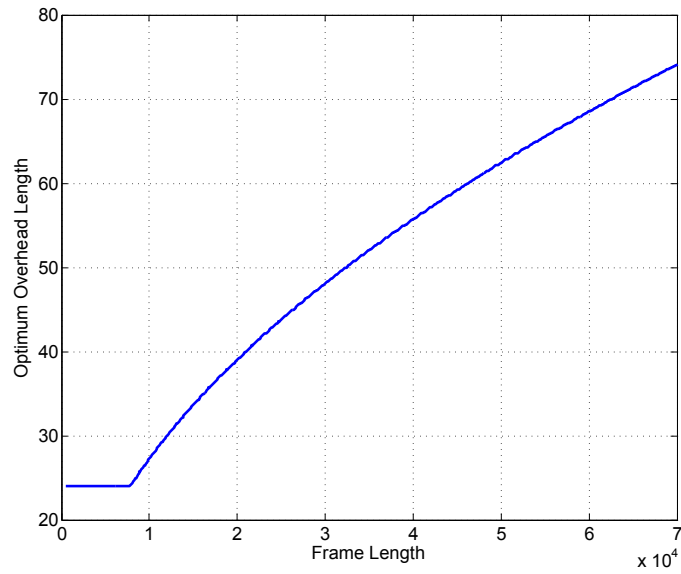


Fig. 2. Optimal Overhead Length vs. Frame Length for a 3 user  $2 \times 2$  system with  $d_i = 1 \forall i$  and  $P_f = P$ .

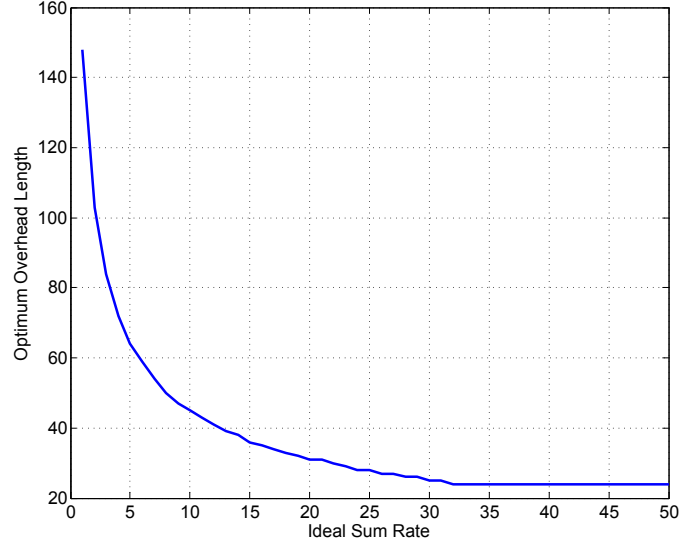


Fig. 3. Optimal Overhead length vs.  $R_{sum}$  for a 3 user  $2 \times 2$  system with  $d_i = 1 \forall i$ , SNR=40dB and  $P_f = P/10$ .

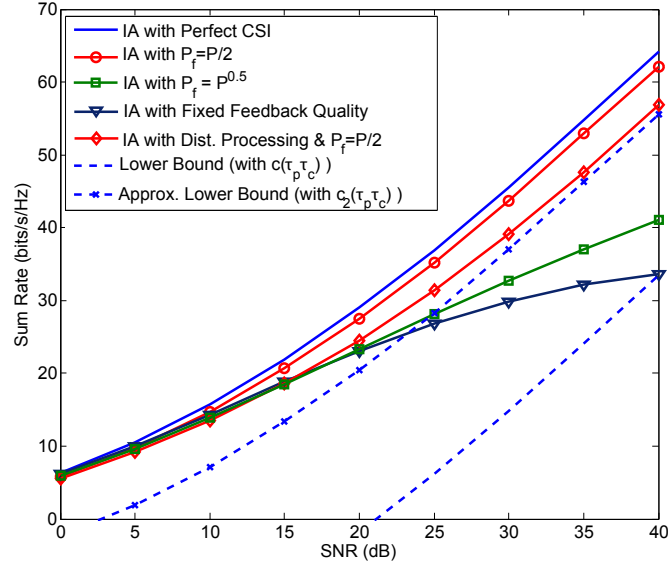


Fig. 4. The sum rate achieved by IA in a 3 users  $5 \times 4$  system ( $d_i = 2 \forall i$ ) with perfect CSI, scaling feedback quality ( $P_f = P/2$ ), slower scaling feedback ( $P_f = \sqrt{P}$ ), and fixed quality feedback with  $SNR_f = 5dB$ .

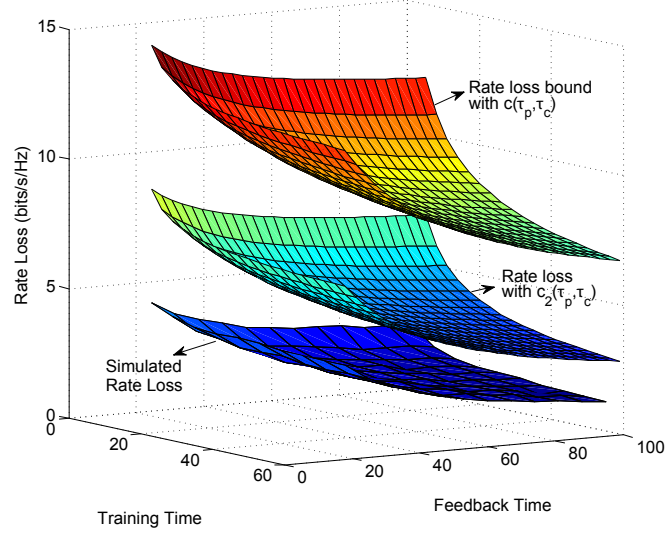


Fig. 5. Rate loss vs. training and feedback times for a 3 user  $2 \times 2$  system with  $d_i = 1 \forall i$  and  $P_f = P/10$ .

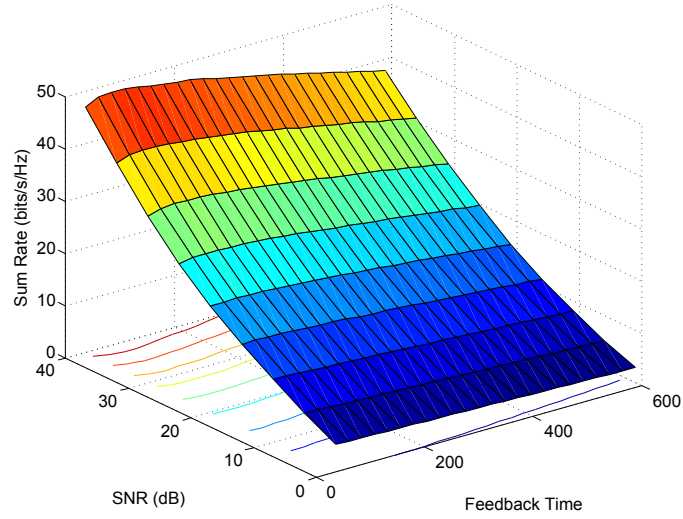


Fig. 6. The effective throughput achieved by interference alignment in a 3 user  $5 \times 4$  with  $P_f = P/100$  and a frame length  $T = 2000$ . This confirms that the amount of feedback needed to achieve optimal throughput is close to minimal.

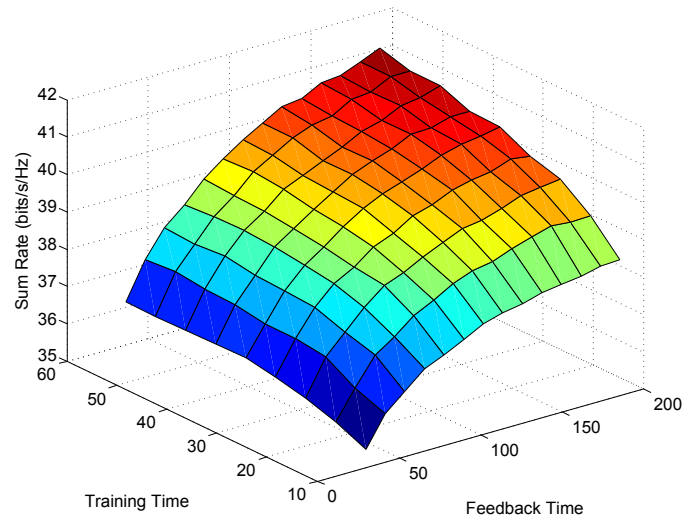


Fig. 7. This plots the sum rate achieved by IA in a 3 user  $5 \times 4$  at 35dB with and feedback SNR of 10dB and a frame length  $T = 10000$ .